

# Failure Detection and Isolation in Integrator Networks

Mohammad Amin Rahimian, Victor M. Preciado \*

**Abstract**—Detection and isolation of link failures under the Laplacian consensus dynamics have been the focus of our previous study. Our results relate the failure of links in the network to jump discontinuities in the derivatives of the output responses of the nodes and exploit that relation to propose failure detection and isolation (FDI) techniques, accordingly. In this work, we extend the results to general linear networked dynamics. In particular, we show that with additional niceties of the integrator networks and the enhanced proofs, we are able to incorporate both unidirectional and bidirectional link failures. At the next step, we extend the available FDI techniques to accommodate the cases of bidirectional link failures and undirected topologies. Computer experiments with large networks and both directed and undirected topologies provide interesting insights as to the role of directionality, as well as the scalability of the proposed FDI techniques with the network size.

## I. INTRODUCTION

Multi-agent network systems have found promising applications in areas such as motion coordination of robots [1]. Cooperative dynamics over a network can be strongly affected by the network failures. Hence, studying the effects of link or node failures on the network dynamics is an important topic in network science and it has various practical implications [2], [3].

Fault Detection and Isolation (FDI) in networked systems is an active area of research with application to power networks [4], and security of cyber-physical systems [5]. In a systematic approach the so-called FDI filter uses available measurements to generate a residual signal, which is then used for fault diagnosis [6]. Various observer and Kalman filter techniques are used to obtain the residual signal [7], which can be then compared with a threshold signal to detect faults. In [8] the residual signal and threshold signals are generated using filtering techniques that allow for noise suppression and less conservative detection thresholds. Failures and attacks are modeled as disturbances in the descriptor systems approach that is adopted in [5], and some fundamental limitations of detection and identification in cyber-physical systems are established from systems-theoretic or graph-theoretic perspectives.

The mathematical framework for investigating the ability to detect and distinguish failures based on the observed output responses is established in [9], where conditions in terms of the inter-nodal distances to the observation points are provided for the detectability of links. Subsequent results in [10] provide a method for detection and isolation of failures under the Laplacian network dynamics. These results relate the presence of discontinuities in the derivatives of particular orders in

the output responses of a subset of nodes to occurrence and location of link failures; and coupled with efficient sensor placement algorithms they can be used to pin down the location of failures in the network. In this paper, we analyze the particular case of networked single integrator dynamics and extend the proofs and FDI algorithms to allow for efficient sensor location in a (directed or undirected) integrator network where link failures can be either unidirectional or bidirectional.

The remainder of this paper is organized as follows. In Section II we set up the notation and give the preliminaries on weighted digraphs and their algebraic properties. Our main analytical findings are set forth in Section III where we first prove the results relating link failures to jump discontinuities of the output derivatives for single-integrator networks, and then consider the case when the links can fail simultaneously in both forward and reverse directions. In Section IV we flesh out the necessary formalism to accommodate both cases of unidirectional and bidirectional link failures within the framework of the previously developed sensor placement algorithms. Computer experiments on large networks in Section V elucidate the results and offer interesting insights on the effects of the network size and edge directionality. Section VI concludes the paper.

## II. PRELIMINARIES

Throughout the paper,  $\emptyset$  is the empty set,  $\mathbb{N}$  denotes the set of all natural numbers, and  $\mathbb{R}$  denotes the set of all real numbers. Also, the set of integers  $\{1, 2, \dots, k\}$  is denoted by  $\mathbb{N}_k$ , and any other set is represented by a calligraphic capital letter. The cardinality of a set  $\mathcal{X}$  is denoted by  $|\mathcal{X}|$ , and  $\mathcal{P}(\mathcal{X}) = \{\mathcal{M}; \mathcal{M} \subset \mathcal{X}\}$  denotes the power-set of  $\mathcal{X}$ , which is the set of all its subsets. The difference of two sets  $\mathcal{X}$  and  $\mathcal{Y}$  is denoted by  $\mathcal{X} \setminus \mathcal{Y}$  and is defined as  $\{x; x \in \mathcal{X} \wedge x \notin \mathcal{Y}\}$ , where  $\wedge$  is the logical conjunction. In addition the logical implication and bi-implication are denoted by  $\rightarrow$  and  $\leftrightarrow$ , respectively, and  $\vee$  denote the logical disjunction. The fixed integer  $n$  represents the number of nodes in the network. Matrices are represented by capital letters, vectors are expressed by boldface lower-case letters, and the superscript  $T$  denotes the matrix transpose. Moreover,  $I$  denotes the identity matrix with proper dimension,  $\mathbf{e}_i$  is the  $i$ -th unit vector in the standard basis of  $\mathbb{R}^n$ , and  $m_{ij} := [M]_{ij}$  indicate the element of matrix  $M$  which is located at its  $i$ -th row and  $j$ -th column.

A directed graph or *digraph* is defined as an ordered pair of sets  $\mathcal{G} := (\mathcal{V}, \mathcal{E})$ , where  $\mathcal{V} = \{\nu_1, \dots, \nu_n\}$  is a set of  $n = |\mathcal{V}|$  vertices and  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  is a set of directed edges. In the graphical representations, each edge  $\epsilon := (\tau, \nu) \in \mathcal{E}$  is depicted by a directed arc from vertex  $\tau \in \mathcal{V}$  to vertex  $\nu \in \mathcal{V}$ . Vertices  $\nu$  and  $\tau$  are referred to as the *head* and *tail* of the edge  $\epsilon$  and a  $(\nu, \nu)$  edge is called a self-loop on  $\nu$ . Given an integer  $k \in \mathbb{N}$ , a set

\* The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104-6228 USA. (email: preciado@seas.upenn.edu).

of (possibly repeated) indices  $\{\alpha_1, \alpha_2, \dots, \alpha_k\} \subseteq \mathbb{N}_n$  and two vertices  $\tau, \nu \in \mathcal{V}$ , an ordered sequence of edges of the form  $\mathcal{W} := (\tau, \nu_{\alpha_1}), (\nu_{\alpha_1}, \nu_{\alpha_2}), \dots, (\nu_{\alpha_{k-1}}, \nu_{\alpha_k}), (\nu_{\alpha_k}, \nu)$  is called a  $\tau\nu$  walk with start-node  $\tau$ , end-node  $\nu$  and length  $k + 1$ . A cycle on node  $\nu$  signifies a  $\nu\nu$  walk. For any  $\{q, p\} \subset \mathbb{N}_n$ ,  $\Omega^k(\mathcal{G}; \nu_q, \nu_p)$  is the set of all  $\nu_q\nu_p$  walks in  $\mathcal{G}$  with length  $k$ . Similarly for  $\{q, r, p\} \subset \mathbb{N}_n$ ,  $\Omega^k(\mathcal{G}; \nu_q, \nu_r, \nu_p) = \{\mathcal{W} \in \Omega^k(\mathcal{G}; \nu_q, \nu_p); (\nu_q, \nu_r) \in \mathcal{W}\}$ , i.e. the set of  $\nu_q\nu_p$  walks with length  $k$  that include the edge  $(\nu_q, \nu_r)$ . In the same venue, the integer  $\text{dist}(\mathcal{G}; \nu_q, \nu_p) = \min\{k \in \mathbb{N} : \Omega^k(\mathcal{G}; \nu_q, \nu_p) \neq \emptyset\}$  is referred to as the distance from  $\nu_q$  to  $\nu_p$  in  $\mathcal{G}$ , and by convention,  $\text{dist}(\mathcal{G}; \nu_q, \nu_q) = 0, \forall q \in \mathbb{N}_n$  and  $\text{dist}(\mathcal{G}; \nu_q, \nu_p) = \infty$  if  $\forall k \in \mathbb{N}, \Omega^k(\mathcal{G}; \nu_q, \nu_p) = \emptyset$ . For ease of notation we usually use  $\text{dist}(\nu_q, \nu_p) := \text{dist}(\mathcal{G}; \nu_q, \nu_p)$ . The diameter of  $\mathcal{G}$ , denoted by  $\text{diam}(\mathcal{G})$ , is defined as the maximum distance between any pair of nodes  $\nu_p, \nu_q \in \mathcal{V}$ :  $\text{diam}(\mathcal{G}) = \max_{\nu_q, \nu_p \in \mathcal{V}} \text{dist}(\nu_q, \nu_p)$ . A matrix  $A \in \mathbb{R}^{n \times n}$  is called an in-weighting on  $\mathcal{G}$  if  $\forall \{\nu_p, \nu_q\} \subset \mathcal{V}, (\nu_q, \nu_p) \notin \mathcal{E} \rightarrow a_{pq} = 0$ . If  $A$  is an in-weighting on  $\mathcal{G}$ , then  $\omega(\mathcal{W}, A) = \prod_{(\nu_q, \nu_p) \in \mathcal{W}} a_{pq}$  is referred to as the weight of walk  $\mathcal{W}$  w.r.t.  $A$ . Given a set of walks  $\Omega$  in digraph  $\mathcal{G}$  and the in-weighting  $A$  on  $\mathcal{G}$ , the function  $\Phi(\Omega, A) = \sum_{\mathcal{W} \in \Omega} \omega(\mathcal{W}, A)$  is defined, which finds use in the proof of the main results in Section III. Note that this function satisfies  $\Phi(\Omega^k(\mathcal{G}; \nu_q, \nu_r, \nu_p), A) = a_{rq} \Phi(\Omega^{k-1}(\mathcal{G}; \nu_r, \nu_p), A)$ . It is also known that given an in-weighting  $A$  on  $\mathcal{G}$  and vertices  $\{\nu_q, \nu_p\} \subset \mathcal{V}$ , that [11], [12]:

$$\Phi(\Omega^k(\mathcal{G}; \nu_q, \nu_p), A) = [A^k]_{pq}. \quad (1)$$

The following definition comes handy in the statement and proof of the main theorem in the next section. At each time  $t_f$  and for all  $p \in \mathbb{N}_n$  and  $k \in \mathbb{N}_z$ , define  $\Delta_{t_f} : \mathbb{N}_n \times \mathbb{N}_z \rightarrow \mathbb{R}$  as  $\Delta_{t_f}(p, k) := \frac{d^k}{dt^k}(x_p)(t_f^+) - \frac{d^k}{dt^k}(x_p)(t_f^-)$ . The function  $\Delta_{t_f}(p, k)$  as defined, measures the jump in the  $k$ -th derivative of the response of agent  $p$  at the time of failure  $t_f$ , and the parameter  $z \leq \text{diam}(\mathcal{G}) + 1$  is a fixed integer, denoting the highest order of derivatives to which the designer has access.

### III. FAILURES IN NETWORKS OF SINGLE-INTEGRATORS

Let us consider a network of  $n$  single-integrators, where each integrator is described by a single state  $x_i$ , with the following dynamics:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{w}(t), \quad t > t_0, \quad (2)$$

where  $\mathbf{x}(t) = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ ,  $\mathbf{w}(t) \in \mathbb{R}^m$ ,  $B \in \mathbb{R}^{n \times m}$ , and  $A \in \mathbb{R}^{n \times n}$  is an in-weighting of graph  $\mathcal{G}$ . We assume that the entries of  $\mathbf{w}(t)$  are  $(\text{diam}(\mathcal{G}) + 1)$ -differentiable. Let us assume link  $\bar{e} = (\nu_j, \nu_i)$  fails at time  $t = t_f$ , resulting in a faulty connectivity graph  $\bar{\mathcal{G}} = (\mathcal{V}, \mathcal{E} \setminus \{\bar{e}\})$  for  $t > t_f$ . The corresponding in-weighting of  $\bar{\mathcal{G}}$ , denoted by  $\bar{A}$ , is a perturbed version of  $A$  that satisfies  $\bar{a}_{ij} = 0$ , and  $\bar{a}_{qr} = a_{qr}$  for all  $q \neq i$ ; i.e. entries on the  $i$ -th row of  $A$  are allowed to change while every other entry remains unaffected.

The following theorem characterizes the effect of link failures on the output derivatives of a network of single integrators:

**Theorem 1.** Consider the dynamic network of single integrators in (2) with output equation  $y_p(t) = x_p(t)$ , and assume link  $\bar{e} = (\nu_j, \nu_i)$  fails at time  $t_f$ . Then, the output of node  $\nu_p$  satisfies:

$$\Delta_{t_f}(p, k) = \Phi(\Omega^{k-1}(\nu_i, \nu_p), A) \sum_{q=1}^N (\bar{a}_{iq} - a_{iq}) x_q(t_f), \quad (3)$$

for  $k = \text{dist}(\nu_j, \nu_p)$ ; and  $\Delta_{t_f}(p, k) = 0$ , for  $k < \text{dist}(\nu_j, \nu_p)$ .

*Proof:* See the Appendix. ■

The next lemma and the following corollary show in particular how the result of Theorem 1 agrees with and maps to the result of Theorem 1 in [13], where we consider the case of networked LTI agents. However, the proofs in [13] rely on Laplace domain techniques rather than the combinatorial arguments in the time-domains that helped us prove Theorem 1; indeed, the flexibilities in the latter case facilitate some extensions that are of particular interest to FDI scenarios.

**Lemma 1.** Consider the failure of link  $\bar{e} = (\nu_j, \nu_i)$  and suppose that  $k = \text{dist}(\nu_j, \nu_p)$ . If the failure of link  $\bar{e}$  is to induce a jump discontinuity in the  $k$ -th derivative of the output response of node  $\nu_p$ , then it should be true that  $\text{dist}(\nu_i, \nu_p) = k - 1 < \text{dist}(\nu_j, \nu_p) = k$ .

*Proof:* See the Appendix. ■

The following is now immediate upon combining the results of Theorem 1 and Lemma 1.

**Corollary 1.** Under the assumptions of Theorem 1,

$$\Delta_{t_f}(p, k) = \begin{cases} c_{i,j,p}, & \text{for } k = (\text{dist}(\nu_i, \nu_p) + 1), \\ 0, & \text{for } k < (\text{dist}(\nu_i, \nu_p) + 1), \end{cases} \quad (4)$$

where  $c_{i,j,p} = \Phi(\Omega^{k-1}(\nu_i, \nu_p), A) \sum_{q=1}^N (\bar{a}_{iq} - a_{iq}) x_q(t_f)$ .

The condition stated in Lemma 1 is intuitive because  $\text{dist}(\nu_i, \nu_p) = k - 1 < \text{dist}(\nu_j, \nu_p) = k$  holds true only if there exist a shortest path of length  $k$  connecting  $\nu_j$  to  $\nu_p$  and with  $(\nu_j, \nu_i)$  as its first edge. In other words, the failed link  $(\nu_j, \nu_i)$  contributes to the flow of information from  $\nu_j$  to  $\nu_i$  as an element of a shortest path from node  $\nu_j$  to node  $\nu_i$ . These observations are in perfect agreement with the sufficient conditions previously studied in [9]. On the other hand, Corollary 1 shows how Theorem 1 may follow as a special case of Theorem 1 in [13] (up to a known constant multiplier), after setting the relative degree  $r = 1$  for the involved networked LTI systems. The proofs in the single-integrator case, however, admit additional niceties that we discuss next.

#### A. Bidirectional Link Failures

To begin, note that the perturbed matrix  $\bar{A}$  is not constrained in the way its entries on the  $i$ -th row are modified, thence Theorem 1 continues to hold in the case where all or several of the edges incoming to node  $\nu_i$  are lost simultaneously. Indeed, it is perceivable for the faults in an agent's hardware or internal structure to cause the failure of multiple links which are incoming to that agent. In the particular case of

a faulty agent  $\nu_i$ , which loses all its incoming links at the instant of failure  $t = t_f$ , the system's dynamics for  $t > t_f$  is characterized by  $\bar{a}_{iq} = 0, \forall q \in \mathbb{N}_n \setminus \{i\}$ . Hence, as a special case, Theorem 1 and the corresponding FDI techniques that are developed in Section IV may also be applied for the detection and isolation of single agent failures by mapping the isolated edges to their head vertices.

On the other hand, for certain applications, where communications are of a bidirectional nature, it is reasonable to consider link failures that simultaneously prevent either agents from communication in the other's direction. Such a failure corresponds to the simultaneous elimination of both links  $\bar{e}$ , defined earlier, and  $\hat{e} = (\nu_i, \nu_j)$  leading to  $\hat{\mathcal{G}} = (\mathcal{V}, \mathcal{E} \setminus \{\bar{e}, \hat{e}\})$  as the information flow structure for  $t > t_f$ . It is worth highlighting that undirected networks, where  $\forall \{\tau, \nu\} \subset \mathcal{V}, (\tau, \nu) \in \mathcal{E} \leftrightarrow (\nu, \tau) \in \mathcal{E}$ , signify the special case that all, not just some, of the links are bidirectional. The FDI methods in this paper are designed to handle the cases where some of the links in the networks are bidirectional and the rest are unidirectional. It is worth pinpointing that, as an assumption of modeling each link is considered either bidirectional or unidirectional, but not both. In other words, if  $(\tau, \nu) \in \mathcal{E}$  and  $(\nu, \tau) \in \mathcal{E}$  and the link between the nodes  $\tau$  and  $\nu$  is specified as bidirectional, then  $(\tau, \nu)$  fails if, and only if,  $(\nu, \tau)$  fails; otherwise, the two links  $(\tau, \nu)$  and  $(\nu, \tau)$  are regarded as separate, and their failures as independent events. Accordingly, it is assumed that the set of all bidirectional links  $\mathcal{B}$  in the network, is known to the designer beforehand.

After modeling the failure of a bidirectional link as the simultaneous failures of two directed links,  $\bar{e} = (\nu_j, \nu_i)$  and  $\hat{e} = (\nu_i, \nu_j)$ , the proof of Theorem 1 can be adapted to yield:

**Proposition 1.** *In the case of the simultaneous failure of  $\bar{e} = (\nu_j, \nu_i)$  and  $\hat{e} = (\nu_i, \nu_j)$ , (4) still holds true if we substitute  $\text{dist}(\nu_i, \nu_p) + 1$  by  $\max\{\text{dist}(\nu_j, \nu_p) + 1, \text{dist}(\nu_i, \nu_p) + 1\}$ .*

*Proof:* See the Appendix. ■

We end this section by an intuitive remark that as each agent of the network system in (2) is a single-integrator, a jump discontinuity (because of a sudden network failure) at point  $\nu_i$  will appear to point  $\nu_p$  after several (length of path) steps of “integrations”. Thus, an agent at point  $\nu_p$  needs to make the same number of “differentiations” before observing the jump due to the failure at point  $\nu_i$ . In what follows we shall see how to determine the observation points along with the required number of differentiations at each point so that the occurrence and location of failures are always inferable from the observed jumped discontinuities.

#### IV. SENSOR PLACEMENT FOR UNIDIRECTIONAL AND BIDIRECTIONAL LINKS

It is assumed that at each instant of time, the designer is given access to the response of a subset of agents, as well as the nominal network information flow digraph  $\mathcal{G}$  (prior to the link failure) and the set of bidirectional  $\mathcal{B}$ . Neither the location of the failure (nodes  $\nu_i$  and  $\nu_j$ ), nor the time of failure  $t_f$  are known to the designer. In the case of detection,

the designer is interested in determining the existence of any single link failure in the network at the instant of failure. For the isolation problem, however, the designer would like to determine “instantaneously”, not only the existence of a failure, but also its location. That is to determine which link, if any, has failed and exactly at the same instant as it fails. The significance of “instantaneous” detection and isolation is better understood upon noting that if the time of failure is random and has a continuous sample space, then “simultaneous” failure of more than one link is a measure zero event; hence, justifying the focus of investigation in this paper, which is on the “single” (possibly bidirectional) link failures. Before shifting attention to the sensor placement problem, two assumptions are set forth:

**Assumption 1.** *For all pairs of nodes  $\nu_p, \nu_q \in \mathcal{V}$ , the in-weighting  $A$  on digraph  $\mathcal{G}$  satisfies  $\Phi(\Omega^{\text{dist}(\nu_q, \nu_p)}(\nu_q, \nu_p), A) \neq 0$ , i.e., the sum of the weights of all shortest paths between them is nonzero.*

**Assumption 2.** *Given the in-weightings  $A$  and  $\bar{A}$  of the faultless and the faulty network,  $\mathcal{G}$  and  $\bar{\mathcal{G}}$ , respectively, we have that  $\sum_{q=1}^N (\bar{a}_{iq} - a_{iq}) x_q(t_f) \neq 0$ , where  $\nu_i$  is the head (or tail too, if the link is bidirectional) of the failed link and  $t_f$  denotes the instant of failure.*

The first assumption above is a provision of consistency that is assumed with regard to the in-weighting matrix  $A$ . This assumption is satisfied almost surely for any assignment of weights on the graph. In particular, it holds true for the Laplacian consensus networks considered in [10]. The second assumption involves the values of the agents' states at the time failure  $t_f$ . This condition also holds true, almost surely, for any in-weightings  $A$ , its perturbed version  $\bar{A}$ , and a random time of failure  $t_f > t_0$ ; since  $\sum_{q=1}^N (\bar{a}_{iq} - a_{iq}) x_q(t_f) = 0$  specifies a low-dimensional hyperplane in the agents' state space that the agents almost surely avoid given a random time of failure.

To enable the designer to handle the desired FDI tasks, she is given access to the output response of a subset of nodes as well as their derivatives up to the  $z$ -th order. In this section, we offer efficient procedures for determining such a subset of nodes, given the network topology and parameter  $z$ ; and in such a way that all link failures in the networks can be detected or isolated from the occurrence of jump discontinuities in the observed outputs and their derivatives. Furthermore, we would like to achieve this goal using as few observation points (sensors) as possible. From Corollary 1 it follows that if the existence of a jump discontinuity in the  $k$ -th derivative of the output response of agent  $p$  is to serve as the basis for a method to detect the failure of edge  $\bar{e}$  at time  $t_f$ , then it should be true that  $\text{dist}(\mathcal{G}; \nu_i, \nu_p) = k - 1 < \text{dist}(\mathcal{G}; \nu_j, \nu_p) = k$ . The latter happens only if there exist a shortest path of length  $k$  connecting  $\nu_j$  to  $\nu_p$  and with  $(\nu_j, \nu_i)$  as its first edge. In [10] we use this observation in the case of Laplacian network dynamics to define binary relations  $\mathcal{R}_k, k \in \mathbb{N}_z$  and  $\mathcal{R}_0$  between the sets  $\mathcal{V}$  and  $\mathcal{E}$  such that for all  $p \in \mathcal{V}$  and

$\epsilon \in \mathcal{E}$  if  $(p, \epsilon) \in \mathcal{R}_k$ , then the failure of link  $\epsilon$  produces a jump in the  $k$ -th derivative of the response of node  $p$  and if  $(p, \epsilon) \in \mathcal{R}_0$  then the failure of edge  $\epsilon$  does not produce a jump in any of the derivatives of the response of node  $p$  upto the  $z$ -th order. We now go ahead and redefine the binary relations per Proposition 1 to accommodate bidirectional link failures. Indeed, bidirectional links are treated specially as any of the two edges in reverse directions can provide us with the required relation for detection when a bidirectional failure occurs.

**Definition 1.** We define the binary relations  $\mathcal{R}_0$  and  $\mathcal{R}_k$  for  $k \in \mathbb{N}_z$ , between  $\mathcal{V}$  and  $\mathcal{E}$ , as follows. For all  $p \in \mathcal{V}$  and  $\epsilon = (\nu_q, \nu_r) \in \mathcal{E}$ , we have that:

- If  $\epsilon \notin \mathcal{B}$ , then  $(p, \epsilon) \in \mathcal{R}_k$  if, and only if,  $\text{dist}(\nu_q, \nu_p) = k$  and  $\text{dist}(\nu_r, \nu_p) = k - 1$ .
- If  $\epsilon \in \mathcal{B}$ , then  $(p, \epsilon) \in \mathcal{R}_k$  if, and only if, one of the following conditions is satisfied:
  - $\text{dist}(\nu_q, \nu_p) = k$  and  $\text{dist}(\nu_r, \nu_p) = k - 1$ , or
  - $\text{dist}(\nu_r, \nu_p) = k$  and  $\text{dist}(\nu_q, \nu_p) = k - 1$ .

The FDI problems can now be posed as follows.

**Problem 1 (Detection).** Given a digraph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , find a subset of nodes  $\mathcal{M}_D \subseteq \mathcal{V}$  of minimum cardinality  $|\mathcal{M}_D|$ , such that for all  $\epsilon \in \mathcal{E}$ , there exists a node  $p \in \mathcal{M}_D$  such that  $(p, \epsilon) \notin \mathcal{R}_0$ .

**Problem 2 (Isolation).** Given a digraph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , find a subset of vertices  $\mathcal{M}_I \subseteq \mathcal{V}$  with the smallest cardinality  $|\mathcal{M}_I|$ , such that  $f_I(\mathcal{M}_I) = 0$ .

The idea for proposing efficient sensor placement algorithms that approximate the solutions of the above problems is by counting the number of edges that are not yet detectable or isolatable from the currently chosen nodes and add new nodes to the existing sets in a greedy manner: in each addition of a new node to the existing sensor set, we aim to decrease the number of edges that are not yet detectable or isolatable as much as possible. To this end, we define a correspondence  $\mathcal{I} : \mathcal{P}(\mathcal{V}) \times \mathcal{E} \rightarrow \mathcal{P}((\mathbb{N}_z \cup \{0\}) \times \mathcal{V})$  which maps an order pair  $(\mathcal{M}, \epsilon)$ , comprised of a sensor set  $\mathcal{M}$  and an edge  $\epsilon$ , to the set of ordered pairs that specify the relations between edge  $\epsilon$  and nodes in  $\mathcal{M}$ . Accordingly, those edges  $\epsilon_1$  and  $\epsilon_2$  which produce the exact same pattern of jumps and at the exact same order of derivatives in the output responses of the nodes in  $\mathcal{M}$  would satisfy  $\mathcal{I}(\mathcal{M}, \epsilon_1) = \mathcal{I}(\mathcal{M}, \epsilon_2)$ ; and none of them can be identified using just the nodes in  $\mathcal{M}$ . We further define two set functions  $f_D$  and  $f_I$  which take a subset of nodes  $\mathcal{M}$  and map it to the number of edges that are, respectively, not detectable or not isolatable using the sensor set  $\mathcal{M}$ . In [13] these functions are shown to be supermodular; wherefore per the theory of submodular set coverings [14], adding nodes greedily with respect to these functions would guarantee that the chosen sensor set is within a factor  $\log(|\mathcal{E}|)$  of the minimal sensor sets that achieve detection or isolation goals (solutions of Problems 1 and 2). The following algorithms are proposed in [10], and included below for completeness, to implement

this idea of supermodular greedy minimization.

---

**Routine 1** Determine a Solution  $\mathcal{M}_D$  to Problem 1

---

**Input:**  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$

- 1:  $\mathcal{M}_D \leftarrow \emptyset$
- 2: **while**  $f_D(\mathcal{M}_D) \neq 0$  **do**
- 3:    $\nu_q \leftarrow \arg \min\{f_D(\mathcal{M}_D \cup \{\nu_q\}) - f_D(\mathcal{M}_D); \nu_q \in \mathcal{V} \setminus \mathcal{M}_D\}$
- 4:    $\mathcal{M}_D \leftarrow \mathcal{M}_D \cup \{\nu_q\}$
- 5: **end while**

**Output:**  $\mathcal{M}_D$

---



---

**Routine 2** Determine a Solution  $\mathcal{M}_I$  to Problem 2

---

**Input:**  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  &  $\mathcal{M}_D$

- 1:  $\mathcal{M}_I \leftarrow \mathcal{M}_D$
- 2: **while**  $f_I(\mathcal{M}_I) \neq 0$  &  $\mathcal{M}_I \neq \mathcal{V}$  **do**
- 3:    $\nu_q \leftarrow \arg \min\{f_I(\mathcal{M}_I \cup \{\nu_q\}) - f_I(\mathcal{M}_I); \nu_q \in \mathcal{V} \setminus \mathcal{M}_I\}$
- 4:    $\mathcal{M}_I \leftarrow \mathcal{M}_I \cup \{\nu_q\}$
- 5: **end while**
- 6: **if**  $f_I(\mathcal{M}_I) \neq 0$  **then**
- 7:    $\mathcal{M}_I \leftarrow \emptyset$
- 8: **end if**

**Output:**  $\mathcal{M}_I$

---

It was noted in Subsection III-A that the set of bidirectional links  $\mathcal{B} \subset \mathcal{E}$  should be made known to the designer. To facilitate the application of Algorithms 1 and 2 to the case of bidirectional link failures in networked single-integrator agents, we define an equivalence relation  $\sim$  on the set  $\mathcal{E}$  that identifies two parallel edges in reverse directions only if they are bidirectional. Specifically, for any  $\{\tau, \nu\} \subset \mathcal{V}$  such that  $(\tau, \nu) \in \mathcal{B}$  and  $(\nu, \tau) \in \mathcal{B}$ , set  $(\tau, \nu) \sim (\nu, \tau)$ , while for any two edges  $\{\alpha, \beta\} \subset \mathcal{E}$  that  $\{\alpha, \beta\} \cap \mathcal{E} \setminus \mathcal{B} \neq \emptyset$ ,  $\alpha \sim \beta$  iff  $\alpha = \beta$ . The task of the equivalence relation  $\sim$  is to identify those edges who share the same head and tail but at opposite directions, only if they are bidirectional. Every other edge in the network is distinguished and therefore identified only with itself. With the afore-defined equivalence relation  $\sim$ , for any edge  $\epsilon \in \mathcal{E}$ ,  $[\epsilon] = \{\hat{\epsilon} \in \mathcal{E}; \hat{\epsilon} \sim \epsilon\}$  denotes the equivalence class of  $\epsilon$ , and for any subset of edges  $\mathcal{X} \subset \mathcal{E}$ ,  $(\mathcal{X} / \sim)$  is the quotient of  $\mathcal{X}$  by  $\sim$ , which is the set of all equivalence classes of the elements of  $\mathcal{X}$ . Last but not least, is the issue of self-loops which are specific to the case of single-integrator agents. In particular, every self-loop would always satisfy an  $\mathcal{R}_0$  relation with all nodes in the network and the proposed algorithms cannot be applied for the detection and isolation of a self-loop  $(\nu_i, \nu_i)$ , although its value (weight) is allowed to change with the failure of a link incoming to node  $\nu_i$ . In the sequel, the set of all self-loops in  $\mathcal{G}$  is denoted by  $\mathcal{H}$ . Next, changing the definitions of the correspondence  $\mathcal{I}(\cdot)$ , and the supermodular functions  $f_D(\cdot)$  and  $f_I(\cdot)$  as follows, allows us to apply Algorithms 1 and 2 to single-integrator networks, while properly identifying bidirectional links and accommodating self-loops. Define for all  $\mathcal{M} \subset \mathcal{V}$  and any of

the equivalence classes in  $\mathcal{E} \setminus \mathcal{H} / \sim$ :

$$\begin{aligned}
f_D : \mathcal{P}(\mathcal{V}) &\rightarrow \mathbb{N}_{|\mathcal{E} \setminus \mathcal{H}|} \cup \{0\}, \\
f_D(\mathcal{M}) &= |\{\iota \in (\mathcal{E} \setminus \mathcal{H} / \sim) : \forall p \in \mathcal{M}, \forall \epsilon \in \iota, (p, \epsilon) \in \mathcal{R}_0\}|, \\
\mathcal{I} : \mathcal{P}(\mathcal{V}) \times (\mathcal{E} \setminus \mathcal{H} / \sim) &\rightarrow \mathcal{P}((\mathbb{N}_z \cup \{0\}) \times \mathcal{V}), \\
\mathcal{I}(\mathcal{M}, [\epsilon]) &= \{(k, p) \in (\mathbb{N}_z \cup \{0\}) \times \mathcal{M} : (p, \epsilon) \in \mathcal{R}_k\}, \\
f_I : \mathcal{P}(\mathcal{V}) &\rightarrow \mathbb{N}_{|\mathcal{E} \setminus \mathcal{H}|} \cup \{0\}, f_I(\mathcal{M}) = \\
|\{\iota \in (\mathcal{E} \setminus \mathcal{H} / \sim) : \exists \hat{\epsilon} \in \mathcal{E} \setminus (\mathcal{H}), \hat{\epsilon} \notin \iota, \mathcal{I}(\mathcal{M}, \iota) &= \mathcal{I}(\mathcal{M}, [\hat{\epsilon}])\}|.
\end{aligned}$$

## V. COMPUTER EXPERIMENTS WITH LARGE NETWORKS

In the following subsections, the performance of the developed routines is tested for different random graph models and with varying model parameters.

### A. A Random Geometric Graph

In a random geometric graph model the nodes of the network are randomly and uniformly spread across a bounded region, and there is an undirected edge between a pair of nodes, wherever a certain distance threshold is met. The graph of Fig. 1(a) depicts one such graph instance with 50 nodes and 200 undirected edges, which are interpreted as pairs of bidirectional links. For this graph a total of nine nodes is sufficient for complete detection, whereas even with all of the nodes observed none of the bidirectional links can be isolated. In other words, for any bidirectional link in the network, there exists at least one other link whose removal will induce the same set of jumps in the entire node-set of the network.

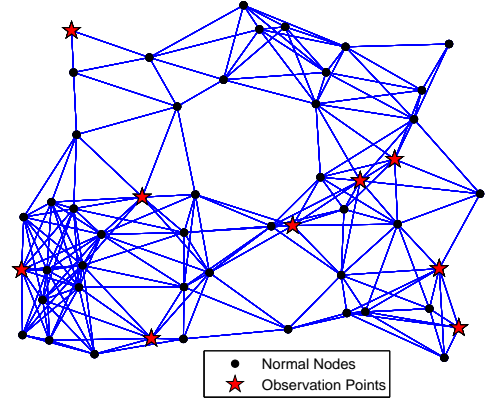
The situation is rather different if the 200 undirected edges of the network in Fig. 1(a) are regarded as 400 unidirectional links. Then the output of Routine 1 has 22 nodes that are indicated in Fig. 1(b), and by observing them the designer can isolate 280 edges out of the total 400. Observing all of the nodes in the network decreases the cardinality of the set of unresolved edges from 120 to just 93, out of the total 400. It is worth highlighting that with the change in the interpretation of the links from bidirectional to unidirectional, matrix  $R$  of the network remains the same, and so does the required highest order of derivatives  $z = 7$ .

Next, each of the undirected edges in Fig. 1(a) is oriented randomly leading to a total of 200 unidirectional edges in Fig. 2. In the latter, a total of 17 nodes is sufficient for detection, and these 17 nodes enable the isolation of all but 75 edges of the digraph, which are highlighted in Fig. 3. For this directed network, by observing all of the nodes in the network, the cardinality of the set of unresolved links reduces to 34.

The preceding results suggest that while detection is achievable more easily in undirected networks, the increased diversity brought about by the directionality of the links improves the isolation task for the case of directed networks.

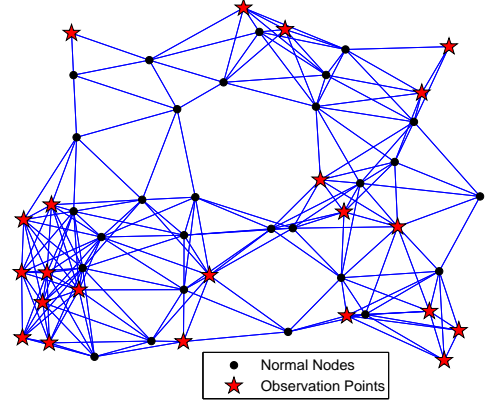
The focus of investigation in the following subsections is shifted to the Erdős-Rényi random graph model, for which the role of edge probability and graph size on the cardinality of the detection set and the highest order of derivatives required is explored.

Choice of Observation Points for Detection in a Random Geometric Graph



(a) Bidirectional Links

Random Geometric Graph with Unidirectional Links



(b) Pairs of Unidirectional Links in Opposite Directions

Fig. 1. (a) With nine observation points and by observing order of derivatives upto  $z = 7$ , every link in this network is detectable, but none of them can be isolated. (b) When the undirected links are regarded as pairs of unidirectional links, the required number of observation points increases to 22 nodes.

### B. Erdős-Rényi Random Graphs: Directed versus Undirected

In a Erdős-Rényi random graph model every potential edge is either existent or not with a fixed probability  $p$ , and independently of all the rest. This model is implemented for varying network sizes  $n$ , and different edge probabilities  $p$ . In Figs.4 and 5, the cardinality of the detection sets  $|\mathcal{M}_D|$  in several randomly generated instances are recorded, averaged, and plotted. The sample means in each case are computed over 50 random instances and the error bars indicate the sample standard deviations for those instances. The plots in all cases confirm the increased difficulty of the detection process for the case of directed networks. Moreover, the cardinality of the detection sets does not scale fast with the network size; an observation which is of practical significance for large networks and complements the theoretical guarantees that are available from the submodular set covering literature. In the case of edge probabilities, however, it is observed that as the edge probabilities approach 1 leading to a complete graph, the

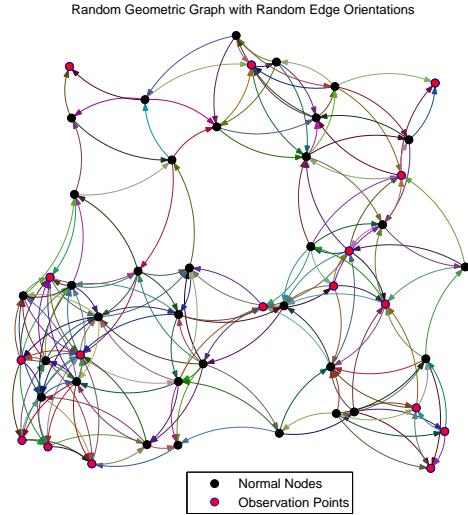


Fig. 2. If the edges of the undirected random geometric graph are randomly oriented, then 17 nodes would be enough to achieve the detection task. This also increases the required highest order of derivatives from  $z = 7$  to  $z = 9$ .

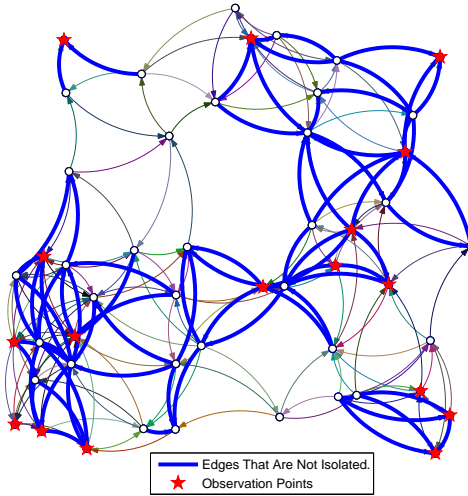


Fig. 3. The 75 edges that are highlighted cannot be isolated using the indicated observation points. However, even with all of the nodes observed, there still remain 34 edges that cannot be isolated.

number of nodes required for detection becomes increasingly large. It is worth highlighting that although for small  $p$  the networks are sparse and can have large diameters, as the edge probability is increased beyond 0.3 the network diameter remains constant at 2 so that only the first three derivatives of any chosen sensor set need to be observed. Similarly for  $n$ , as the network size is increased beyond 60, the network diameters remain fixed at 3 and only the first four derivatives of the outputs in any chosen sensor set are observed.

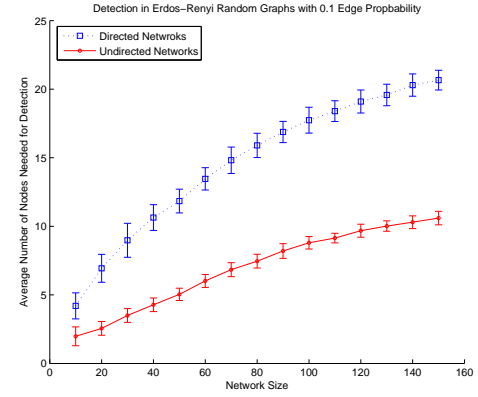


Fig. 4.  $|\mathcal{M}_D|$  versus  $n$  for Erdős-Rényi random graphs with  $p = 0.1$

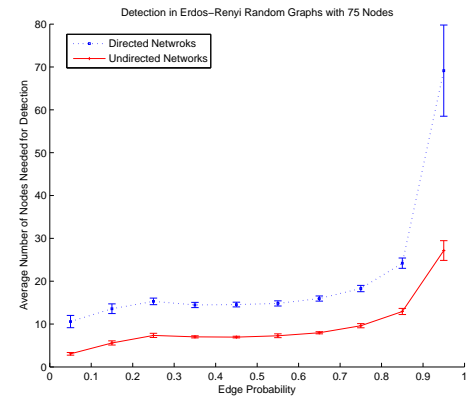


Fig. 5.  $|\mathcal{M}_D|$  versus  $p$  for Erdős-Rényi random graphs with  $n = 75$

## VI. CONCLUSIONS

In this paper, we developed FDI techniques for single-integrator networks that enable the designer to detect and isolate link failures based on the observed jumps in the derivatives of the output responses of a subset of nodes by relating the jumps in the derivatives at the time of failure to the distance of the failed link from the observation point. Our results covered both cases of unidirectional and bidirectional link failures. We also extended our previously developed sensor placement algorithms to accommodate both types of link failures (unidirectional and bidirectional). These algorithms were tested in large random networks, and the results suggest that link failures in directed networks are harder to detect but easier to isolate, as compared to undirected networks. The latter effect can be attributed to the increased diversity that is brought about by the directionality of the links. Moreover, both the cardinality of the detection sets and the required order of derivatives are shown to scale up reasonably well with the network size, and this agrees with the performance guarantees that are available from the theory of submodular set coverings and bound the size of the chosen sets to within a multiplicative  $\log(|\mathcal{E}|)$  factor of the minimal sensor set, where  $|\mathcal{E}|$  is the size of the edge set.

APPENDIX  
PROOFS OF THE MAIN RESULTS

A. Theorem 1

Given an initial condition  $\mathbf{x}_0 := \mathbf{x}(t_0) \in \mathbb{R}^N$ , the solution to (2) for  $t_0 \leq t \leq t_f$  is trivially given by:

$$\mathbf{x}(t) = e^{A(t-t_0)} \mathbf{x}_0 + \int_{t_0}^t e^{A(t-\tau)} B \mathbf{w}(\tau) d\tau, \text{ for } t_f \geq t \geq t_0. \quad (5)$$

The evolution of states after failure is therefore governed by the state matrix  $\bar{A}$ , instead of  $A$ , as follows

$$\mathbf{x}(t) = \left( e^{\bar{A}(t-t_f)} \mathbf{x}_f + \int_{t_f}^t e^{\bar{A}(t-\tau)} B \mathbf{w}(\tau) d\tau \right), \text{ for } t \geq t_f, \quad (6)$$

where  $\mathbf{x}_f := \mathbf{x}(t_f^-)$ , i.e., the state of the faultless evolution at the instant right before failure.

For any  $p \in \mathbb{N}_N$  fixed, differentiating (5) and (6)  $k$  times and using the Leibniz integral rule yields for  $t_f > t > t_0$  that:

$$\begin{aligned} \frac{d^k}{dt^k} x_p(t) = & \quad (7) \\ \mathbf{e}_p^T \left( A^k \mathbf{x}(t) + \sum_{m=0}^{k-2} A^m B \frac{d^{k-m-1}}{dt^{k-m-1}} \mathbf{w}(t) + A^{k-1} B \mathbf{w}(t) \right), \end{aligned}$$

and for  $t > t_f$  that,

$$\begin{aligned} \frac{d^k}{dt^k} x_p(t) = & \quad (8) \\ \mathbf{e}_p^T \left( \bar{A}^k \mathbf{x}(t) + \sum_{m=0}^{k-2} \bar{A}^m B \frac{d^{k-m-1}}{dt^{k-m-1}} \mathbf{w}(t) + \bar{A}^{k-1} B \mathbf{w}(t) \right). \end{aligned}$$

Next, note that by differentiability of  $\mathbf{w}(t)$  and continuity of the states,  $\frac{d^k}{dt^k} \mathbf{w}(t)|_{t=t_f^+} = \frac{d^k}{dt^k} \mathbf{w}(t)|_{t=t_f^-} = \frac{d^k}{dt^k} \mathbf{w}(t)|_{t=t_f}, \forall k \in \mathbb{N}$  and  $\mathbf{x}(t_f^+) = \mathbf{x}(t_f^-)$ . Hence, subtracting the two equations in (7) and (8) for  $t_f^+$  and  $t_f^-$  yields:

$$\begin{aligned} \Delta_{t_f}(p, k) = & \mathbf{e}_p^T (\bar{A}^k - A^k) \mathbf{x}(t_f) + \\ & \mathbf{e}_p^T (\bar{A}^{k-1} - A^{k-1}) B \mathbf{w}(t_f) + \\ & \mathbf{e}_p^T \sum_{m=0}^{k-2} (\bar{A}^m - A^m) B \left( \frac{d^{k-m-1}}{dt^{k-m-1}} \mathbf{w}(t) \Big|_{t=t_f} \right), \quad (9) \end{aligned}$$

With  $p, \bar{A}$  and  $A$  fixed in the preceding and  $k \in \mathbb{N}_{\text{dist}(\nu_j, \nu_p)}$ , for all  $m \in \mathbb{N}_{\text{dist}(\nu_j, \nu_p)}$ , define  $\nabla(q, m) = [\bar{A}^m]_{pq} - [A^m]_{pq}$ , so that (9) can be rewritten as:

$$\begin{aligned} \Delta_{t_f}(p, k) = & \sum_{q=1}^N \nabla(q, k) x_q(t_f) + \sum_{q=1}^N \nabla(q, k-1) [B \mathbf{w}(t_f)]_q \\ & + \sum_{m=0}^{k-2} \sum_{q=1}^N \nabla(q, m) \left[ B \frac{d^{k-m-1}}{dt^{k-m-1}} \mathbf{w}(t_f) \right]_q \quad (10) \end{aligned}$$

Next to compute  $\nabla(q, m)$ , substitute for  $[A^m]_{pq}$  and  $[\bar{A}^m]_{pq}$

from (1) to get:

$$\nabla(q, m) = \Phi(\Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_p), \bar{A}) - \Phi(\Omega^m(\mathcal{G}; \nu_q, \nu_p), A).$$

By partitioning the sets  $\Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_p)$  and  $\Omega^m(\mathcal{G}; \nu_q, \nu_p)$ , (A) can be rewritten as:

$$\begin{aligned} \nabla(q, m) = & [\Phi(\Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_p) \setminus \Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_i, \nu_p), \bar{A}) \\ & + \Phi(\Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_i, \nu_p), \bar{A})] - \\ & [\Phi(\Omega^m(\mathcal{G}; \nu_q, \nu_p) \setminus \Omega^m(\mathcal{G}; \nu_q, \nu_i, \nu_p), A) \\ & + \Phi(\Omega^m(\mathcal{G}; \nu_q, \nu_i, \nu_p), A)]. \quad (11) \end{aligned}$$

Next note that none of the walks in  $\Omega^m(\mathcal{G}; \nu_q, \nu_p) \setminus \Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_i, \nu_p)$  or  $\Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_p) \setminus \Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_i, \nu_p)$  can include  $(\nu_j, \nu_i)$  as an edge for any  $j \in \mathbb{N}_N$ . This is true, since otherwise if there exists a walk  $\mathcal{W}_1$  that violates the above, then removing the segment of  $\mathcal{W}_1$  from  $\nu_q$  to  $\nu_i$ , which consists of at least two edges, one to reach  $\nu_j$  from  $\nu_q$  followed by the edge  $(\nu_j, \nu_i)$ , yields a new  $\nu_i \nu_p$  walk  $\mathcal{W}_2$  with length at most  $m-2$ . Now  $(\nu_j, \nu_i) \mathcal{W}_2$  is a  $\nu_j \nu_p$  walk of length at most  $m-1$  in  $\mathcal{G}$ , which is a contradiction, since  $m \leq \text{dist}(\nu_j, \nu_p)$ . It next follows that  $\Phi(\Omega^m(\mathcal{G}; \nu_q, \nu_p) \setminus \Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_i, \nu_p), A) = \Phi(\Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_p) \setminus \Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_i, \nu_p), \bar{A})$ , as none of the walks involved include any of the edges  $(\nu_q, \nu_i)$  for  $q \in \mathbb{N}_N$ , and these are the only edges at which the digraphs  $\bar{\mathcal{G}}$  and  $\mathcal{G}$  or the in-weightings  $\bar{A}$  and  $A$  differ. Hence, (A) simplifies into:

$$\begin{aligned} \nabla(q, m) = & \quad (12) \\ \Phi(\Omega^m(\bar{\mathcal{G}}; \nu_q, \nu_i, \nu_p), \bar{A}) - \Phi(\Omega^m(\mathcal{G}; \nu_q, \nu_i, \nu_p), A) \\ = & \bar{a}_{iq} \Phi(\Omega^{m-1}(\bar{\mathcal{G}}; \nu_i, \nu_p), \bar{A}) - a_{iq} \Phi(\Omega^{m-1}(\mathcal{G}; \nu_i, \nu_p), A). \end{aligned}$$

The last step in deriving a simplified expression for  $\nabla(q, m)$  is to argue that  $\Phi(\Omega^{m-1}(\mathcal{G}; \nu_i, \nu_p), A) = \Phi(\Omega^{m-1}(\bar{\mathcal{G}}; \nu_i, \nu_p), \bar{A})$ . To see why, note that since  $\bar{\mathcal{G}}$  is derived upon removal of the edge  $(\nu_j, \nu_i)$  from  $\mathcal{G}$ , it follows that

$$\text{dist}(\nu_i, \nu_p) = \text{dist}(\bar{\mathcal{G}}; \nu_i, \nu_p) \geq \text{dist}(\nu_j, \nu_p) - 1, \quad (13)$$

where the digraph argument for the distance function indicate that the distances are calculated with respect to the edge-removed digraph  $\bar{\mathcal{G}}$  as opposed to usual case where the distances are calculated with respect to the original digraph  $\mathcal{G}$ . The inequalities in (13) together with  $\text{dist}(\nu_j, \nu_p) \geq m$ , implies that  $\text{dist}(\nu_i, \nu_p) = \text{dist}(\bar{\mathcal{G}}; \nu_i, \nu_p) \geq m-1$ , so that none of the walks in  $\Omega^{m-1}(\mathcal{G}; \nu_i, \nu_p)$  or  $\Omega^{m-1}(\bar{\mathcal{G}}; \nu_i, \nu_p)$  can include any  $(\nu_q, \nu_i)$  edges,  $\forall q \in \mathbb{N}_N$ , and  $\Phi(\Omega^{m-1}(\mathcal{G}; \nu_i, \nu_p), A) = \Phi(\Omega^{m-1}(\bar{\mathcal{G}}; \nu_i, \nu_p), \bar{A})$ ; hence,

$$\nabla(q, m) = (\bar{a}_{iq} - a_{iq}) \Phi(\Omega^{m-1}(\mathcal{G}; \nu_i, \nu_p), A), \quad (14)$$



which upon replacement in (10) yields:

$$\begin{aligned} \Delta_{t_f}(p, k) &= \Phi(\Omega^{k-1}(\mathcal{G}; \nu_i, \nu_p), A) \sum_{q=1}^N (\bar{a}_{iq} - a_{iq}) x_q(t_f) + \\ &\Phi(\Omega^{k-2}(\mathcal{G}; \nu_i, \nu_p), A) \sum_{q=1}^N (\bar{a}_{iq} - a_{iq}) [B\mathbf{w}(t_f)]_q \\ &+ \sum_{m=1}^{k-2} \Phi(\Omega^{m-1}(\mathcal{G}; \nu_i, \nu_p), A) \sum_{q=1}^N (\bar{a}_{iq} - a_{iq}) \\ &\left[ B \frac{d^{k-m-1}}{dt^{k-m-1}} \mathbf{w}(t_f) \right]_q. \end{aligned} \quad (15)$$

To complete the proof, note that for  $m \leq \text{dist}(\nu_j, \nu_p) - 1$ ,  $\Omega^{m-1}(\mathcal{G}; \nu_i, \nu_p) = \emptyset$ , since for any  $\mathcal{W}_3 \in \Omega^{m-1}(\mathcal{G}; \nu_i, \nu_p)$ ,  $(\nu_j, \nu_i)\mathcal{W}_3$  is a  $\nu_j\nu_p$  walk of length  $m$  which contradicts with  $m \leq \text{dist}(\nu_j, \nu_i) - 1$ . Hence, for  $k \leq \text{dist}(\nu_j, \nu_p)$ , (15) simplifies into  $\Delta_{t_f}(p, k) = \Phi(\Omega^{k-1}(\mathcal{G}; \nu_i, \nu_p), A) \sum_{q=1}^N (\bar{a}_{iq} - a_{iq}) x_q(t_f)$ , thus completing the proof for the case of the failure of the single link  $\bar{e}$ . That for  $k < \text{dist}(\nu_j, \nu_p)$ ,  $\Delta_{t_f}(p, k) = 0$  also follows as  $\Omega^{k-1}(\mathcal{G}; \nu_i, \nu_p) = \emptyset$  for any such  $k$ . ■

#### B. Lemma 1

Notice that any shortest path from  $\nu_i$  to  $\nu_p$  of length  $k - 1$  gives a path of length  $k$  from  $\nu_j$  to  $\nu_p$ , whence given  $k = \text{dist}(\nu_j, \nu_p)$ , it follows that:

$$\text{dist}(\nu_i, \nu_p) \geq k - 1. \quad (16)$$

Also notice that if  $\text{dist}(\nu_i, \nu_p) \leq k - 1$  there is no path of length  $k - 1$  from  $\nu_i$  to  $\nu_p$ , i.e.

$$\begin{aligned} \Phi(\Omega^{k-1}(\mathcal{G}; \nu_i, \nu_p), A) &\text{ is nonzero,} \\ \text{only if } \text{dist}(\nu_i, \nu_p) &\leq k - 1. \end{aligned} \quad (17)$$

Now (16) and (17) together imply that for the right-hand side of (3) in Theorem 1 to be non-zero when  $k = \text{dist}(\nu_j, \nu_p)$ , it should be true that  $\text{dist}(\nu_i, \nu_p) + 1 = k$ , which is the same as the claimed condition. ■

#### C. Proposition 1

To see how Theorem 1 applies to the case of bidirectional link failures, in parallelism with  $\bar{A}$  and  $\bar{\mathcal{G}}$  in the preceding, let  $\hat{A} \neq A$  be an in-weighting on  $\hat{\mathcal{G}}$  that denotes the perturbed version of  $A$  following the simultaneous failure of links  $\bar{e}$  and  $\hat{e}$ . The perturbations only affect the entries of  $A$  on its  $i$ -th and  $j$ -th rows, such that  $\hat{a}_{ij} = \hat{a}_{ji} = 0$ , while  $\hat{a}_{qr} = a_{qr}$ ,  $\forall r \in \mathbb{N}_N$  and  $\forall q \in \mathbb{N}_N \setminus \{i, j\}$ . For any agent  $p \in \mathbb{N}_N$  and  $t \geq t_f$  the evolution of the state of agent  $p$  following the simultaneous failure of links  $\bar{e}$  and  $\hat{e}$  is given by (6) with  $\hat{A}$  substituted for  $\bar{A}$ . Repeating the same procedure as in the proof of Theorem 1 leads to the proof of the proposition as follows.

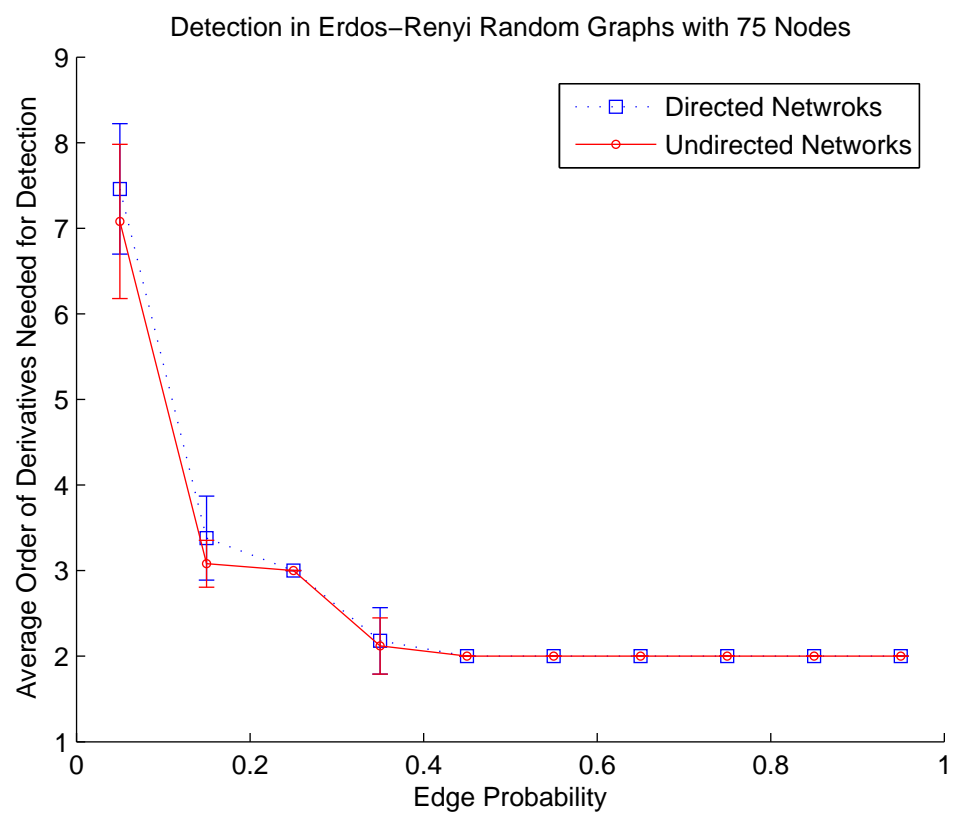
For the case of simultaneous failure of  $\bar{e}$  and  $\hat{e}$ , note that (7) to (11) continue to hold after substituting  $\hat{A}$  and  $\hat{\mathcal{G}}$  for  $\bar{A}$  and  $\bar{\mathcal{G}}$ , respectively. The transitions from (11) to (12) and (14) also carry through with the same replacements and upon the additional observation that the walks in  $\Omega^m(\mathcal{G}; \nu_q, \nu_p) \setminus \Omega^m(\mathcal{G}; \nu_q, \nu_i, \nu_p)$ ,  $\Omega^m(\hat{\mathcal{G}}; \nu_q, \nu_p) \setminus \Omega^m(\hat{\mathcal{G}}; \nu_q, \nu_i, \nu_p)$ ,  $\Omega^{m-1}(\mathcal{G}; \nu_i, \nu_p)$  and

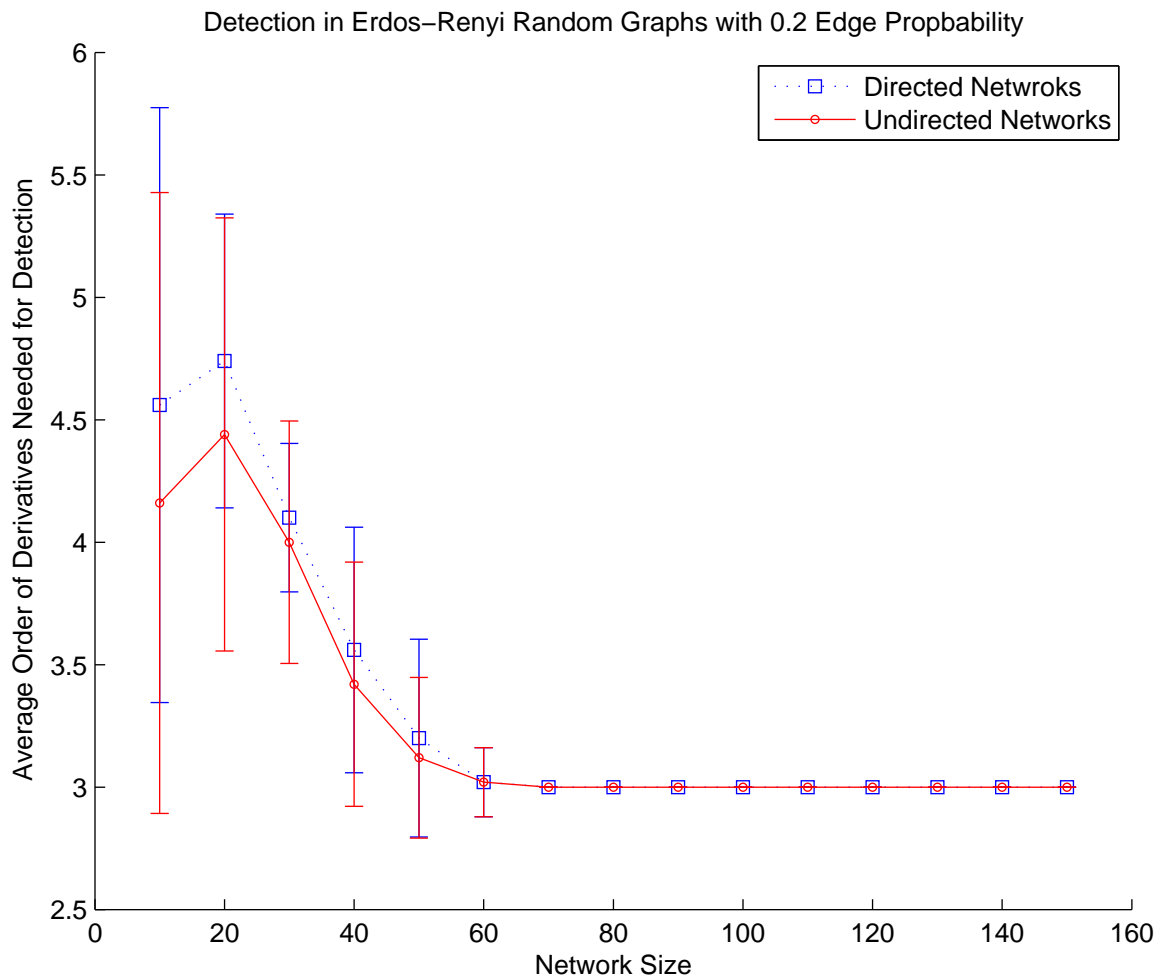
$\Omega^{m-1}(\hat{\mathcal{G}}; \nu_i, \nu_p)$  include neither any  $(\nu_q, \nu_i)$  edges, as stated in the previous case, nor any  $(\nu_q, \nu_j)$  edges,  $\forall q \in \mathbb{N}_N$ . The rest of the proof is identical to the previous case, except for  $\bar{A}$  which should be replaced with  $\hat{A}$ . ■

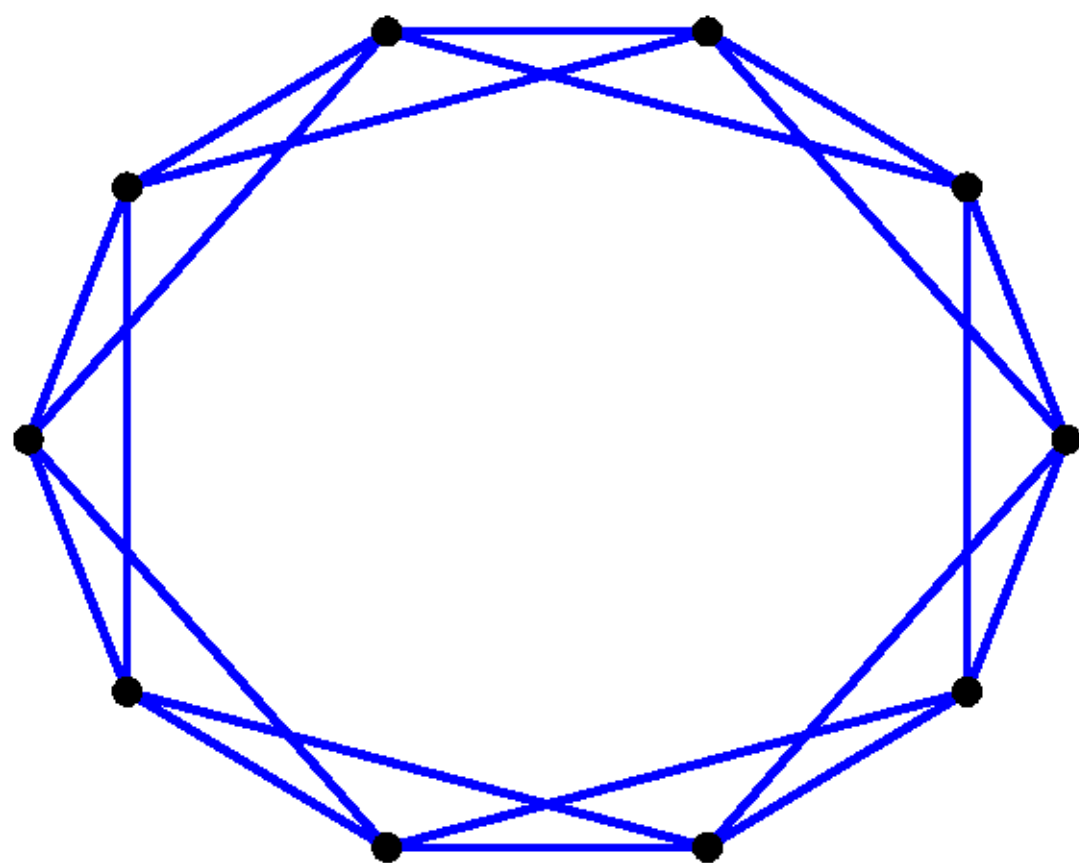
#### REFERENCES

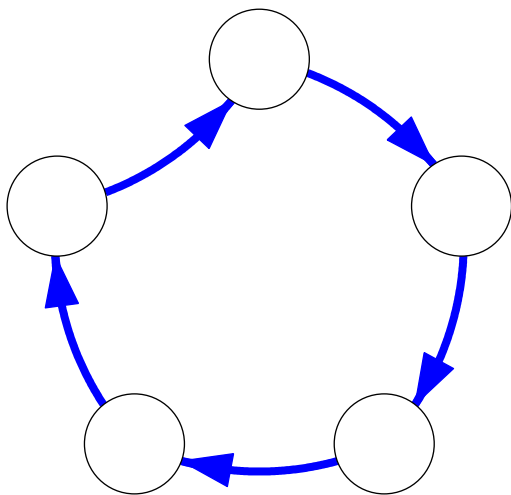
- [1] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*. Princeton University Press, 2010.
- [2] M. A. Rahimian and A. G. Aghdam, "Structural controllability of multi-agent networks: Robustness against simultaneous failures," *Automatica*, vol. 49, no. 11, pp. 3149 – 3157, 2013.
- [3] J. Kleinberg, M. Sandler, and A. Slivkins, "Network failure detection and graph connectivity," *SIAM Journal on Computing*, vol. 38, no. 4, pp. 1330–1346, 2008.
- [4] W. Pan, Y. Yuan, H. Sandberg, J. Gonçalves, and G.-B. Stan, "Real-time fault diagnosis for large-scale nonlinear power networks," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, 2013, pp. 2340–2345.
- [5] F. Pasqualetti, F. Dorfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *Automatic Control, IEEE Transactions on*, vol. 58, no. 11, pp. 2715–2729, 2013.
- [6] E. E. Tiniou, P. M. Esfahani, and J. Lygeros, "Fault detection with discrete-time measurements: An application for the cyber security of power networks," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, 2013, pp. 194–199.
- [7] P. Menon and C. Edwards, "Robust fault estimation using relative information in linear multi-agent networks," *Automatic Control, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2013.
- [8] C. Keliris, M. Polycarpou, and T. Parisini, "A distributed fault detection filtering approach for a class of interconnected continuous-time nonlinear systems," *Automatic Control, IEEE Transactions on*, vol. 58, no. 8, pp. 2032–2047, 2013.
- [9] M. A. Rahimian, A. Ajorlou, and A. G. Aghdam, "Digraphs with distinguishable dynamics under the multi-agent agreement protocol," *Asian Journal of Control*, 2014, in press.
- [10] M. A. Rahimian and V. M. Preciado, "Detection and isolation of link failures under the agreement protocol," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, 2013, pp. 7364–7369.
- [11] N. Biggs, *Algebraic Graph Theory*. Cambridge University Press, 1994.
- [12] V. M. Preciado and A. Jadbabaie, "Moment-based spectral analysis of large-scale networks using local structural information," *IEEE/ACM Transactions on Networking*, vol. 21, no. 2, pp. 373–382, Apr. 2013.
- [13] M. A. Rahimian and V. M. Preciado, "Detection and isolation of failures in directed networks of lti systems," *arXiv preprint arXiv:1408.3164*, 2014.
- [14] L. Wolsey, "An analysis of the greedy algorithm for the submodular set covering problem," *Combinatorica*, vol. 2, no. 4, pp. 385 – 393, 1982.



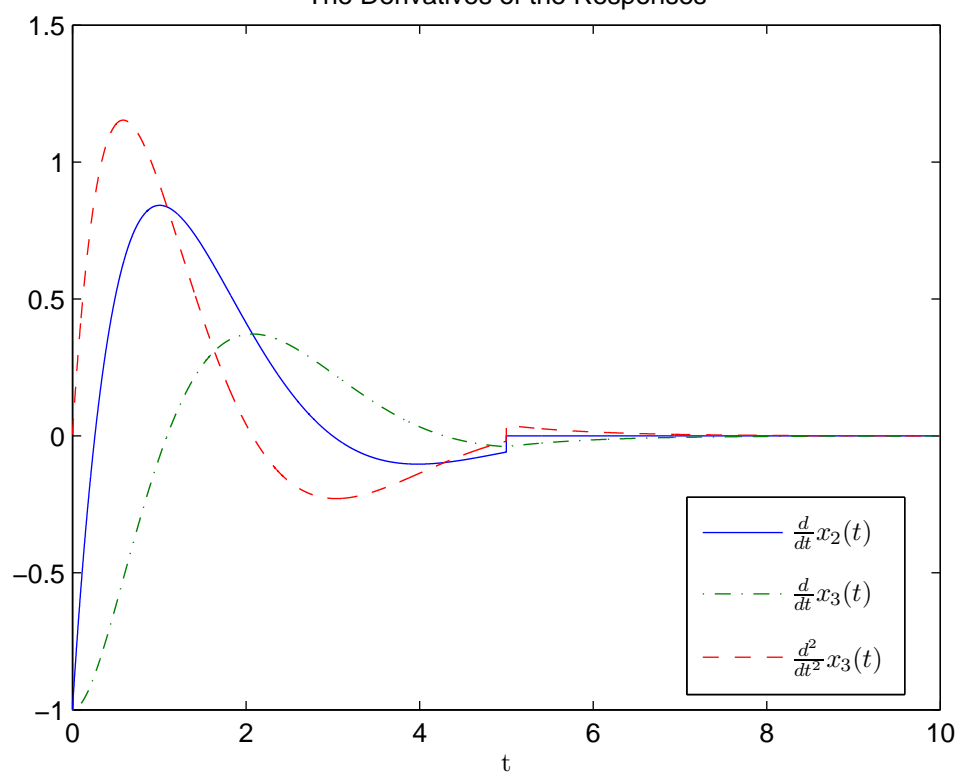


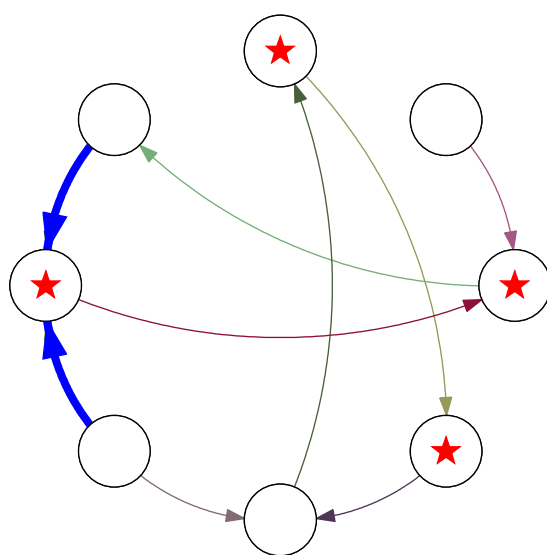
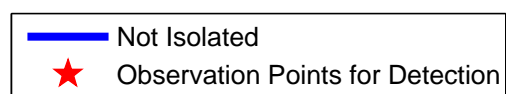




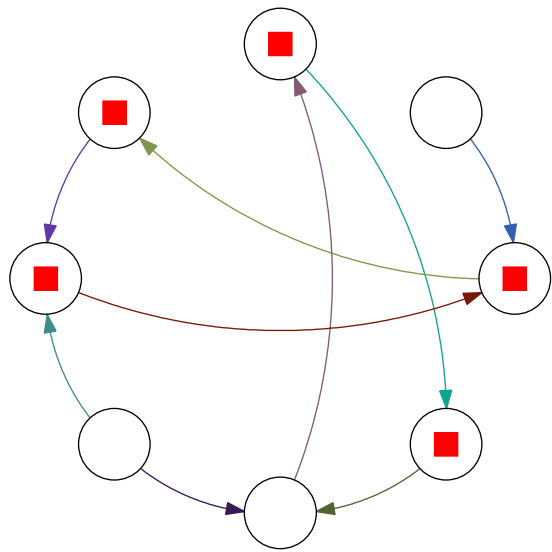


The Derivatives of the Responses

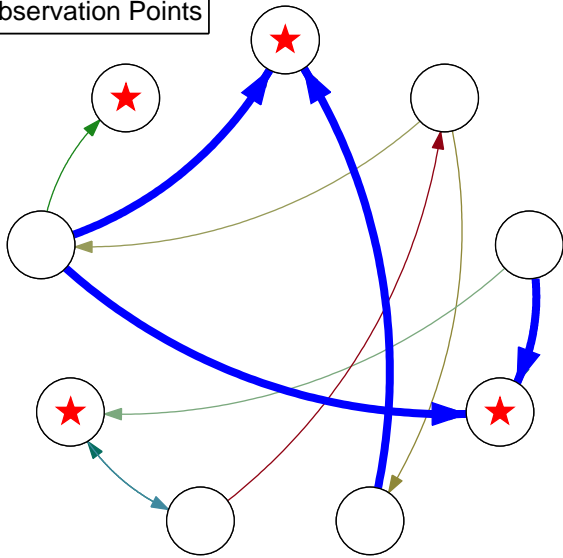
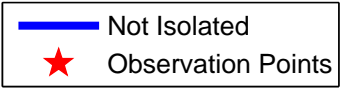




■ Observation Points for Isolation







★ Observation Points

